

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 09:23:09 ; Search time 11.3094 Seconds

(without alignments)
2442.495 Million cell updates/sec

Title: US-09-836-077-3

Perfect score: 3615

Sequence: 1 MTPPPGRAPRSAPRRVPG.....LAASLWGLVPLTGLGLVH 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Archived: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3615	100.0	666	SM7A_HUMAN	075326 homo sapien
2	3214	88.9	664	SM7A_MOUSE	09498 mus musculu
3	654.5	18.1	771	SM3A_HUMAN	014563 homo sapien
4	635	17.6	772	SM3A_MOUSE	008665 mus musculu
5	632.5	17.5	772	SM3A_CHICK	090607 gallus gall
6	628.5	17.4	748	SM3B_MOUSE	062177 mus musculu
7	623	17.2	772	SM3A_RAT	063548 rattus norv
8	621	17.2	860	SZ1A_BRARE	09w7j1 brachydanio
9	617	17.1	749	SM3B_HUMAN	013214 homo sapien
10	617	17.1	782	SM4B_MOUSE	062179 mus musculu
11	613	17.0	778	SZ1B_BRARE	09w686 brachydanio
12	598.5	16.6	861	SM4D_MOUSE	009126 mus musculu
13	597.5	16.5	751	SM3C_MOUSE	062181 mus musculu
14	597	16.5	785	SM3F_HUMAN	013375 homo sapien
15	595.5	16.5	764	SM22_BRARE	09w696 brachydanio
16	591.5	16.4	751	SM3C_HUMAN	099885 homo sapien
17	588.5	16.3	673	SM4B_HUMAN	09np22 homo sapien
18	582	16.1	785	SM3F_MOUSE	088632 mus musculu
19	575	15.9	751	SM3C_CHICK	042236 gallus gall
20	570	15.8	862	SM4D_HUMAN	092854 homo sapien
21	567.5	15.7	775	SM3E_HUMAN	015041 homo sapien
22	565	15.6	761	SM3D_CHICK	090663 gallus gall
23	555	15.4	775	SM3E_MOUSE	095075 mus musculu
24	548.5	15.2	777	SM3D_HUMAN	095025 homo sapien
25	538.5	14.9	785	SM3E_CHICK	042237 gallus gall
26	536	14.8	770	SM4F_HUMAN	095754 homo sapien
27	525.5	14.5	766	SM2F_BRARE	09y1x4 brachydanio
28	515.5	14.3	777	SM2F_MOUSE	09z123 mus musculu
29	510	14.1	776	SM4F_RAT	09z143 rattus norv
30	506.5	14.0	838	SM4G_HUMAN	09n1n9 homo sapien
31	491.5	13.6	837	SM4G_MOUSE	09w1n7 mus musculu
32	485.5	13.4	760	SM4A_MOUSE	062178 mus musculu
33	482.5	13.3	762	SM4A_HUMAN	09h3s1 homo sapien

34	479.5	13.3	834	1	SM4C_MOUSE	064151 mus musculu
35	452	12.5	730	1	SM1A_SCHAM	026473 schistocerc
36	435	12.0	295	1	SM4D_CHICK	090665 gallus gall
37	419	11.6	697	1	SM2A_SCHGR	09x2c8 schistocerc
38	418	11.6	888	1	SM6A_MOUSE	035464 mus musculu
39	411	11.4	712	1	SM1A_CAEL	017330 caenorhabd1
40	409	11.3	1074	1	SM5A_HUMAN	013591 homo sapien
41	408	11.3	712	1	SM1A_HUMAN	026972 tribolium c
42	407	11.3	771	1	SM1A_TRICF	024322 drosophila
43	407	11.3	1093	1	SM5B_MOUSE	060519 mus musculu
44	403	11.1	1077	1	SM5A_MOUSE	062217 mus musculu
45	402	11.1	706	1	SM2A_DROME	024323 drosophila

ALIGNMENTS

RESULT 1	ID	SM7A_HUMAN	STANDARD:	PRT:	666 AA.
AC	075326:				
DT	16-OCT-2001 (rel. 40, Created)				
DT	16-OCT-2001 (rel. 40, Last sequence update)				
DT	16-OCT-2001 (rel. 40, Last annotation update)				
DE	Semaphorin 7A precursor (Semaphorin L) (Sema L) (Semaphorin				
DE	K1) (Sema K1) (John-Milton-Hargen human blood group Ag) (JMH blood				
DE	group antigen) (CD108 antigen) (CDw108).				
GN	SEMA7A OR SEMAL OR CD108.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=98389619; PubMed=9721204;				
RA	Lange C., Liehr T., Goen M., Gebhart E., Fleckenstein B., Ensser A.;				
RT	"New eukaryotic semaphorins with close homology to semaphorins of DNA				
RT	viruses.";				
RL	Genomics 51:340-350(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99218431; PubMed=10201933;				
RA	Yamada A., Kubo K., Takeshita T., Harashima N., Kawano K., Mine T.,				
RA	Sagawa K., Sugamura K., Itoh K.;				
RT	"Molecular cloning of a glycosylphosphatidylinositol-anchored molecule				
RT	CDw108.";				
RL	J. Immunol. 162:4094-4100(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RT	TISSUE=Placenta;				
RX	MEDLINE=98380463; PubMed=9712866;				
RA	Xu X., Ng S., Wu Z.-L., Nguyen D., Homburger S., Seidel-Dugan C.,				
RA	Edens A., Luo Y.;				
RT	"Human semaphorin K1 is glycosylphosphatidylinositol-linked and				
RT	defines a new subfamily of viral-related semaphorins.";				
RL	J. Biol. Chem. 273:22428-22434(1998).				
RN	[4]				
RP	CHARACTERIZATION.				
RX	MEDLINE=99344620; PubMed=10416131;				
RA	Angelisova P., Dibal K., Cerny J., Hliger I., Horejsi V.;				
RT	"Characterization of the human leukocyte GPI-anchored glycoprotein				
RT	CDw108 and its relation to other similar molecules.";				
RL	Immunobiology 200:234-245(1999).				
CC	- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE NERVOUS SYSTEM AND IN				
CC	MODULATING IMMUNE FUNCTION.				
CC	- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.				
CC	- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA, TESTIS, OVARY, SPLEEN,				
CC	BRAIN, SPINAL CHORD, LUNG, HEART, ADRENAL GLAND, LYMPH NODES,				
CC	THYMUS, INTESTINE AND KIDNEY.				
CC	- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.				
CC	- SIMILARITY: CONTAINS 1 SEMA DOMAIN.				
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.				

FT	DOMAIN	642	729	IG-LIKE C2-TYPE DOMAIN.
FT	DISULFID	649	722	ARG-LYS-RICH (BASIC).
FT	CARBOHYD	53	53	BY SIMILARITY
FT	CARBOHYD	125	125	N-LINKED (GLICNAC. . .) (POTENTIAL).
FT	CARBOHYD	590	590	N-LINKED (GLICNAC. . .) (POTENTIAL).
SO	SEQUENCE	771 AA:	88889 MM;	9985F8D3FEAE8456 CRC64;

Query Match	18.1%:	Score 654.5:	DB 1:	Length 771:
Best Local Similarity	28.2%:	Pred. No. 1,1e-41:		
Matches 194:	Conservative 115:	Mismatches 260:	Indels 99:	Gaps 27:

QY	32	RLILLMAAASAGHLNSG----	PRIFAVMKGVGQDRVD----	FGOTEPHYLFHEP	82
Db <td>6</td> <td>RIVCFMGVLLTARANYQGNKNNVRLKSYKEMLESNNVITFNGLANSSSYTFLEDE</td> <td>65</td> <td></td> <td></td>	6	RIVCFMGVLLTARANYQGNKNNVRLKSYKEMLESNNVITFNGLANSSSYTFLEDE	65		
QY <td>83</td> <td>GSSVWVGGRGVVYLFDPREGKNASRVYVNIQSTK-----</td> <td>GSLCDKDCENYITLLER</td> <td>135</td> <td></td>	83	GSSVWVGGRGVVYLFDPREGKNASRVYVNIQSTK-----	GSLCDKDCENYITLLER	135	
Db <td>66</td> <td>RSRLVYGKDHIEFSFDLVNINDFOKIYWPVSYTRDECKMAGKDLK-ECANFIVLVA</td> <td>123</td> <td></td> <td></td>	66	RSRLVYGKDHIEFSFDLVNINDFOKIYWPVSYTRDECKMAGKDLK-ECANFIVLVA	123		
QY <td>136</td> <td>RSE-GLLACGTARAPRSPCWNLVNGTVVPLGEM-----</td> <td>RGVAPFSDENS</td> <td>178</td> <td></td>	136	RSE-GLLACGTARAPRSPCWNLVNGTVVPLGEM-----	RGVAPFSDENS	178	
Db <td>124</td> <td>YNQTHLYACGTAPRPIG-----</td> <td>TYIELGHPEDNIFLENSHFENGKGRPYDPKLLT</td> <td>177</td> <td></td>	124	YNQTHLYACGTAPRPIG-----	TYIELGHPEDNIFLENSHFENGKGRPYDPKLLT	177	
QY <td>179</td> <td>LVLEFGDEVYSTRKQEVNGKIPRRRIIGSEELT---</td> <td>SDTVMONPQITKATVHQ-DQ</td> <td>234</td> <td></td>	179	LVLEFGDEVYSTRKQEVNGKIPRRRIIGSEELT---	SDTVMONPQITKATVHQ-DQ	234	
Db <td>178</td> <td>ASLLIDGELYSCTGAADFMGRDPAIRFTLGHHPIRTEQDSRWLMDPKFISALISESN</td> <td>237</td> <td></td> <td></td>	178	ASLLIDGELYSCTGAADFMGRDPAIRFTLGHHPIRTEQDSRWLMDPKFISALISESN	237		
QY <td>235</td> <td>AYDDKIYFFERDNDPKNEAPLANSRVAYQLCRGDOGESSLSVSKWNTFLKAMLYCSDA</td> <td>294</td> <td></td> <td></td>	235	AYDDKIYFFERDNDPKNEAPLANSRVAYQLCRGDOGESSLSVSKWNTFLKAMLYCSDA	294		
Db <td>238</td> <td>PEDDKVFYFFERNALIDGESHKATARIQIQIQKDNFGHRSI-VKWTITFLKARLLICSP</td> <td>296</td> <td></td> <td></td>	238	PEDDKVFYFFERNALIDGESHKATARIQIQIQKDNFGHRSI-VKWTITFLKARLLICSP	296		
QY <td>295</td> <td>ATN---KMFNRLQDVFL--PDPSQCMRDTRYGVF--</td> <td>SFPMWTSAVCVYSLGDIDKVF</td> <td>346</td> <td></td>	295	ATN---KMFNRLQDVFL--PDPSQCMRDTRYGVF--	SFPMWTSAVCVYSLGDIDKVF	346	
Db <td>297</td> <td>GPNGIDTFHDELDQVFLNMFKDP---</td> <td>KNPVYGVFTTSSNIFKGSAYCVMGMSDVRVF</td> <td>352</td> <td></td>	297	GPNGIDTFHDELDQVFLNMFKDP---	KNPVYGVFTTSSNIFKGSAYCVMGMSDVRVF	352	
QY <td>347</td> <td>-----RTSSLKGVHSLPNPRPGKC-----</td> <td>LPDQDIPETPEQVADRHPVADR</td> <td>390</td> <td></td>	347	-----RTSSLKGVHSLPNPRPGKC-----	LPDQDIPETPEQVADRHPVADR	390	
Db <td>353</td> <td>LGPYAHROGPNQWPPYQGRVYPRPGRGTCSPKSTFGGFSBTKLPDPDYITFAASHPMATNP</td> <td>412</td> <td></td> <td></td>	353	LGPYAHROGPNQWPPYQGRVYPRPGRGTCSPKSTFGGFSBTKLPDPDYITFAASHPMATNP	412		
QY <td>391</td> <td>VEPMG---PLKTPLEHSHKYHYQKVAVHRMQASHETFEVLYLTDRGTIHKVE--</td> <td>PGHQ</td> <td>445</td> <td></td>	391	VEPMG---PLKTPLEHSHKYHYQKVAVHRMQASHETFEVLYLTDRGTIHKVE--	PGHQ	445	
Db <td>413</td> <td>VEPMNRPPIVITDV--NYQFTQIVYDVBDAEDGQ-YDMFICGIDVGTALIKVSSIPKET</td> <td>468</td> <td></td> <td></td>	413	VEPMNRPPIVITDV--NYQFTQIVYDVBDAEDGQ-YDMFICGIDVGTALIKVSSIPKET	468		
QY <td>446</td> <td>EHSFANFM-ELQPRRAALIQTMSLDMERKRLVYSSQMEVSOVPLDCEVYGGGCHGL</td> <td>504</td> <td></td> <td></td>	446	EHSFANFM-ELQPRRAALIQTMSLDMERKRLVYSSQMEVSOVPLDCEVYGGGCHGL	504		
Db <td>469</td> <td>WDLDEVILLEETVTRPEPTAISAMELSTQOQLYGSTAGVACQALPIHRHRCIDYIGKACAECC</td> <td>528</td> <td></td> <td></td>	469	WDLDEVILLEETVTRPEPTAISAMELSTQOQLYGSTAGVACQALPIHRHRCIDYIGKACAECC	528		
QY <td>505</td> <td>MSRDEYCGMDGRCISIV--SSERSVLQSIINAPFHKKECPNPKPDK---</td> <td>APLQVSL-A</td> <td>557</td> <td></td>	505	MSRDEYCGMDGRCISIV--SSERSVLQSIINAPFHKKECPNPKPDK---	APLQVSL-A	557	
Db <td>529</td> <td>LARDPYCAMDGSACGRYPTAKRRTRRDIDIRGDDPLTHCSDLHNDHNHSHSEERLIYGV</td> <td>588</td> <td></td> <td></td>	529	LARDPYCAMDGSACGRYPTAKRRTRRDIDIRGDDPLTHCSDLHNDHNHSHSEERLIYGV	588		
QY <td>558</td> <td>PNRNYLYSCPMESRHATYVSMRHKEVNEQSCERGHOSPNCI---</td> <td>LFLENLTAQOYGYHF</td> <td>612</td> <td></td>	558	PNRNYLYSCPMESRHATYVSMRHKEVNEQSCERGHOSPNCI---	LFLENLTAQOYGYHF	612	
Db <td>589</td> <td>ENSSFTFLCSPSQALYVWQFORNEERKEEIRVDHIITRDQGLLRSLQOKDSGNYL</td> <td>648</td> <td></td> <td></td>	589	ENSSFTFLCSPSQALYVWQFORNEERKEEIRVDHIITRDQGLLRSLQOKDSGNYL	648		
QY <td>613</td> <td>CEAOGSYFREAOHQWLLPEDGIMAEHL</td> <td>640</td> <td></td> <td></td>	613	CEAOGSYFREAOHQWLLPEDGIMAEHL	640		
Db <td>649</td> <td>CHAVEHGFIQTLKVTL---EVIDTEHL</td> <td>673</td> <td></td> <td></td>	649	CHAVEHGFIQTLKVTL---EVIDTEHL	673		

RESULT 4	SM3A.MOUSE	STANDARD:	PRT:	772 AA.
AC	008665:	062180:	062215:	
DT	30-MAY-2000	(Rel. 39,	Created)	
DT	30-MAY-2000	(Rel. 39,	Last sequence update)	
DT	16-OCT-2001	(Rel. 40,	Last annotation update)	
DE	Semaphorin 3A precursor	(Semaphorin III)	(Sema III)	(Semaphorin D)
DE	(Sema D).			
GN	SEM3A OR SEM4D OR SEMD.			
OC	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

OC	Mammalia	Eutheria;	Rodentia;	Sciurognathi;	Muridae;	Murinae;	Mus.
OX	NCBI_TaxID=10090;						
RN	[1]						
RC	SEQUENCE FROM N.A.						
RC	STRAIN=NMRI; TISSUE=Embryo;						
RX	MEDLINE=95267431; PubMed=7748561;						
RA	Pueschel A.W., Adams R.H., Betz H.;						
RT	"Murine semaphorin D/collapsin is a member of a diverse gene family						
RL	and creates domains inhibitory for axonal extension.";						
RM	Neuron 14:941-948(1995).						
RN	[2]						
RA	SEQUENCE FROM N.A.						
RX	MEDLINE=97470885; PubMed=9331345;						
RA	Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,						
RT	Yagi T.;						
RT	"Disruption of semaphorin II/D gene causes severe abnormality in						
RL	peripheral nerve projection.";						
RN	Neuron 19:519-530(1997).						
RM	[3]						
RA	SEQUENCE FROM N.A.						
RA	Kimura T., Fishman M.C.;						
RT	"cDNA sequence of mouse collapsin/semaphorin III.";						
RL	Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.						
RN	[4]						
RC	SEQUENCE OF 107-772 FROM N.A.						
RC	TISSUE=Fetal brain;						
RX	MEDLINE=95267432; PubMed=7748562;						
RA	Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,						
RA	Goodman C.S., Kolodkin A.L.;						
RT	"Semaphorin III can function as a selective chemorepellent to pattern						
RT	sensory projections in the spinal cord.";						
RL	Neuron 14:949-959(1995).						
CC	-I- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO						
CC	PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT						
CC	NORMALLY TERMINATE DORSALLY.						
CC	-I- SUBCELLULAR LOCATION: Secreted.						
CC	-I- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT						
CC	(E11). IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.						
CC	EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.						
CC	-I- DOMAIN: STRONG BINDING TO NEROPILIN IS MEDIATED BY THE CARBOXY						
CC	THIRD OF THE PROTEIN.						
CC	-I- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.						
CC	-I- SIMILARITY: CONTAINS 1 SEMA DOMAIN.						
CC	-I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use by non-profit institutions as long as its content is in no way						
CC	modified and this statement is not removed. Usage by and for commercial						
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/						
CC	or send an email to license@isb-sib.ch).						
CC	-----						
DR	EMBL; X85993; CAAS9985.1; -						
DR	EMBL; D85028; BAA19773.1; -						
DR	EMBL; LA1541; AAL77611.1; -						
DR	EMBL; LA0484; AAA73934.1; -						
DR	MCD; MGI:107558; Sema3a.						
DR	InterPro; IPR003599; Ig.						
DR	InterPro; IPR003006; Ig_MHC.						
DR	InterPro; IPR003659; Plexin-like.						
DR	InterPro; IPR001627; Sema.						
DR	Pfam; PF00047; Ig; 1.						
DR	Pfam; PF01403; Sema; 1.						
DR	SMART; SM00409; Ig; 1.						
DR	SMART; SM00423; PSI; 1.						
KW	Signal; Immunoglobulin domain; Multigene family; Neurogenesis;						
KW	Developmental protein; Glycoprotein.						
FT	SIGNAL	1	20	POTENTIAL.			
FT	CHAIN	21	772	SEMAPHORIN 3A.			
FT	DOMAIN	240	538	SEMA.			
FT	DOMAIN	643	730	IG-LIKE C2-TYPE DOMAIN.			
FT	DOMAIN	728	770	ARG/LYS-RICH (BASIC)..			

```

FT DISULFID 650 723 BY SIMILARITY.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 193 193 D -> N (IN REF. 4).
FT CONFLICT 207 207 H -> D (IN REF. 1).
FT CONFLICT 253 253 D -> G (IN REF. 1).
FT CONFLICT 352 352 F -> L (IN REF. 4).
FT CONFLICT 403 403 A -> G (IN REF. 1).
FT CONFLICT 571 572 OH -> ED (IN REF. 1).
FT CONFLICT 616 620 EDRKE -> RRSKR (IN REF. 1).
FT CONFLICT 623 623 R -> K (IN REF. 4).
SQ SEQUENCE 772 AA; 88799 MW; E89A08528B10AEC3 CRC64;

```

Query Match 17.6%; Score 635; DB 1; Length 772;
 Best Local Similarity 27.9%; Pred. No. 3,3e-40;
 Matches 192; Conservative 109; Mismatches 281; Indels 106; Gaps 27;

```

36 LMAAASAOGLHRSQ---PRIFAVMKGVGQDRVD-----FCOTEPTVLFHEPGSS 86
10 LFMGLTLARANYANGKNNVPRKLSTYKEMLESNNVITFNGLANSSSYHTFLDEE-RSR 68
87 VAVGGRGKYVLPDPFGKNAVRYVNIQSTK-----GSLDKKDCENTITLLERRSE- 138
69 LYVGAKDHIFSFENLVNIKDFOKIWPVSYTRDECKMAGKDIK-ECANFIKYLEAVNOT 127
139 GLLAGCTAARHPSCWNLVNGTVPLGEM-----RGYAFEPDENSELVF 182
128 HLYAGCTGAFHPICT-----TYIEVGHHPEDNIFKLQDSHFENGKSKSPYDKLLASLL 181
183 EGDEVYSTIRKQYNGKIPFRIRIGSELYT---SDTVMONPOPKATIV-HODQAYVD 238
182 IDGELYSGTADEFGMRDAIFRTLGHHNPITREQDSRWLNDRPFIASHLIPESDPED 241
239 KIYTFREDNPKNAPLPLANSRVNQLCRGQGGSSLSVSKWNTFLKMLVCSDAATN- 297
242 KYEFERENAIIDEGHSGKATARIQICNDKDFGHRSL-VNKTFTFLKRLKISVPGPN 300
298 --KNNRRLQDVEFL--PDRSGQWRDTRVYGVF---SNPNYSAVCVSLGDIKDYF--- 346
301 IDTHDELQDVFLLMSKDP-----KNPIYGVTTSSNIFKGSAAVCMTSISDVRVFLGTY 356
347 -----RTSSLLKGYHSSLPRPRGKC-----LPDQDPIPTETFOVADRHPVAVRPEM 394
357 AHRDQPNQWVRYQGRVPRPGTGRPSKTFGGFSDTKLDRDVIITFARSHRAMVNPVFI 416
395 G-----PLKTPLEHSHYHOKAVANHMOKASHGETFENVLYITTRGTTHKAYE--PGQESHF 449
417 NNRPIMIKTDV---NYQFTQIVVDVDAEDGQ-YDVMFTGTVGVTLKVVSVPKETWHD 472
450 AFNIM-ELIOPFRRAAIIQMSLDAERKRLVSSQWEVSQVPLDLCEVVGCGHGLMSRD 508
473 EVELLEBMTVFERPTTISAMELSTKOOLYIGTAGVADLPRLHRCDIYKACAECCCLAD 532
509 PYCGMDGRCISIV--SSERSVLOSINPAERPKECRN-----PKRDKAPLOKVSIA 557
533 PYCADMGSSCSYFPTAKRTRRORDIRNGDPLTHCSDLQHNHNGHRSLEERITVGE-- 590
558 PMSRYYLQPMESHRATYSMRKKEVQSCERGHOSPNQI-----LFIENLTAQOYGHNF 612
591 -NSSTFLECSPKSRQRLVYVQFORNEDKREIIRMGDHIITREGGLILSLRLOKKSQGNL 649
613 CEAGSGYFREAQHWOLLPEGDIMAENL 640
650 CHAVEHGMQTLKYTL---EVIDTEHL 674

```

RESULT 5
 SM3A_CHICK STANDARD; PRT: 772 AA.
 AC 090607;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3A precursor (Collapsin-1) (COLL-1).
GN SEMA3A OR COLL.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
RP 666-680.
RC TISSUE=Brain;
RX MEDLINE=94006554; PubMed=8402908;
RA Luo Y., Raible D., Raper J.A.;
RT "Collapse: a protein in brain that induces the collapse and paralysis
RT of neuronal growth cones."
RL Cell 75:217-227(1993).
CC -1- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROPILIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
CC AND MUSCLE, MODERATE LEVELS IN LIVER, COLLAPLIN-1, -2, -3, AND -5 BIND TO
CC VIRTUALLY ABSENT IN LIVER, COLLAPLIN-1, -2, -3, AND -5 BIND TO
CC OVERLAPPING BUT DISTINCT AXON TRACTS.
CC -1- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: 002528; AAC59638.1; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR pfam: PF00047; Ig_1.
DR pfam: PF01403; Sema; 1.
DR SMART: SM00409; Ig; 1.
DR SMART: SM00423; PSI; 1.
DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
KW developmental protein; Glycoprotein.
KW SIGNAL
FT CHAIN 1 23 772 SEMAPHORIN 3A.
FT DOMAIN 240 538 SEMA.
FT DOMAIN 643 730 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 728 770 ARG/LYS-RICH (BASIS).
FT DISULFID 650 723 BY SIMILARITY.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 772 AA; 88867 MW; E91E09DECC940AC CRC64;

```

Query Match 17.5%; Score 632.5; DB 1; Length 772;
 Best Local Similarity 28.4%; Pred. No. 5e-40;

Matches 195; Conservative 101; Mismatches 265; Indels 125; Gaps 28;

```

29 LRLRLLLMAAASAOGLHRSQPRIFAVMKGVGQDR-VDF---GQTEPTVLFHEPG 83
7 IALLSLGVLLAGVNCQHVKNVPRKLSTYKEMLESNNVITFNGLANSSSYHTFLDEE- 65
84 SSSVWVGGRGKYVLPDPFGKNAVRYVNI-----GSKTKSCL-----DKRDE 127
66 RSRLYVGAKDHIFSFNL-----VNKEYOKIWPVSYTRDECKMAGKDIKLRCA 115

```

```

QY 128 NYITLLERSE-GLACGTNARHPSGMNVLNGTVPLGEM-----RGTA 170
DB 116 NIKLKYNNQTHLACGAGAPMC-----TYLEVGHEDNIFRMDSHFENGKGS 169
QY 171 PESPENSLVLEEGDEVSTIRKOEYNGKIPRRIRIGSELYT---SDTVQNPQFIKA 227
DB 170 PYDPRLLTASLVDDELIVSGTADFMGRDFAIFRLIGHHHPRTDQHSRLMLNDREFISA 229
QY 228 TIV-HODAYVDKIYFFREDNPKNPEAPLNVSRVAOLCRDGGESSLSVSKMNTFLK 286
DB 230 HLPESDNPEDDKIYFFRENALDGEHNGKATHARIGQCKNDGSHSL-VNKKWTFELK 288
QY 287 AMLVCSDAATN---KNENRLOVFL--PDPSGQMDTRIVGVF---SNPMYKSAVCYS 338
DB 289 AKLISVPGPNIDITHFELDQVFLMNSKDP---KNPLIVGVFTTSSIFKGSVAVCMS 344
QY 339 LGDIDKVF-----RTSSLKGYSLLPNRPGKC-----LPDQDPIETEFQVAD 382
DB 345 MTDVARVFLGPRAHADGPNYQVRYQGRVYPRPGTCSKTGDFDSTKDLDEVITFAR 404
QY 383 RHPEVAQRYEPMGP---LKTPLFHSKYHYQKAVAHMQASHGETFHVLYLTTRGTIHK 438
DB 405 SHPAMYNVFPIINSRPIKTKDV---DYQFQIVDVAEDGQ-YDVMFICTDIGTVLK 460
QY 439 VVE-PGEQHSFAFNIM-EIOFPRAAAIQTMSLDAERKLYSSQWESQVPLDCEY 496
DB 461 VVSIKRETHLEELVLEEMTFEPTVISAARKISTKOOILIGSATGVSQPLHRCDY 520
QY 497 GGGCHGCLMSRDPYCGMDOGRCSISY---SSERSVLOSINPAEPHEC-----PNPKPDK 548
DB 521 GKACECLLARDPYCAMGSSCSRYFPKAKRTRQDIRNGDPLRHCSDLQHDNDPSCOT 580
QY 549 APLQKVSILAPNSRYLLSCPMESRHAATYSWR-----HKEVNEOS-----CEPGHQSPPC 596
DB 581 LEEKIYGVENSFTLECSPPKSORAIYVQFOKNDHKVEIKVDRIKRTGOG----- 634
QY 597 ILFENLTAQOYGHCECAOEGSYFR 622
DB 635 -LLKSLDORSDGTYFCHAVEHGTIO 659

RESULT 6
SM3B_MOUSE
ID SM3B_MOUSE STANDARD: PRT: 748 AA.
AC Q62177;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin A) (Sema A).
GN SEMA3B OR SEMA OR SEMA.
OS Mus musculus (Mouse).
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX [1]
RN
RC SEQUENCE FROM N.A.
RA STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; Pubmed=7748561;
RA Pleschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFIC TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DB EMBL; X85990; CAA5982.1; -
DB MGD; MGI:107561; Sema3b.
DB InterPro: IPR003599; 1g.
DB InterPro: IPR003006; 1g.MHC.
DB InterPro: IPR003659; Plexin-like.
DB InterPro: IPR001627; Sema.
DB Pfam; PF00047; 1g.
DB SMART; SM00409; IG; 1.
DB SMART; SM00423; PSI; 1.
KW Signal. Immunoglobulin domain. Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25
FT CHAIN 26 748 SEMAPHORIN 3B.
FT DOMAIN 239 536 SEMA.
FT DOMAIN 636 716 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 698 702 POLY-GLY.
FT DOMAIN 723 743 ARG-RICH (BASIC).
FT DISULF 643 709 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 748 AA; 82894 MW; 1668BD2397C9305 CnC64;

Query Match 17.4%; Score 628.5; DB 1; Length 748;
Best Local Similarity 31.5%; Pred. No. 9.7e-40;
Matches 181; Conservative 76; Mismatches 244; Indels 73; Gaps 21;

QY 125 DEENITLLERSE-GLACGTNARHPSG---WNLVNGTVPLG-----EMRGVAPESP 174
DB 112 EGMNVRLHAYNNHLLACRTGAFHPTCALRMWATAGTASTPEKLEDKGKTPIDP 171
QY 175 DENSVLFEQDEVSTIRKOEYNGKIPRRIRIGSELYT---SDTVQNPQFIKA-TIV 230
DB 172 RHRPSVVLGEELVSGVYADLKGROFTIFRSLGQPSLRTEPHDSRWMLNEPFAVFWNP 231
QY 231 HODAYVDKIYFFREDNPKNPE-APLNVSRVAOLCRDGGESSLSVSKMNTFLKML 289
DB 232 ESENDPDKIYFFRESAVEAAPAMGRMSVSRVQICRNDLGQSHL-VNKKWTFELKARL 290
QY 290 VCS---DAATNKNFNLOVFLLPDPSGQMDT-RYGVFSNP---MNSAVCVSLGID 343
DB 291 VCSYVGECDTHFDOLQVFL---SSRDROTPLLYAVFTSISSGVFGSAYCVISMNDVR 347
QY 344 KVF-----RTSSLKGYSLLPNRPGKC-----LPDQDPIETEFQVADRHPEV 387
DB 348 RAFLGLPLHKRGPTHQWVSQGRVYPRPGMCPSTKFTGFSSTKDFPDVIOGSHNPLM 407
QY 388 AQRVPRPKPLTPLF---HSKYHYQKAVAHMQASHGETFHVLYLTTRGTIHKY--YEP 442
DB 408 YNPVLPNG--GRPLFLQYAGAGYTFQIADRYAADG--YVLFGLGTGVYLVKYSVPK 464
QY 443 GEOSFSPFNIMEIOFPFRAAIIQTMSLDAERKLYSSQWESQVPLDCEVYGGCGH 502
DB 465 GRPNSGELLLEELQVFDSDAITSMTQISSKROULYASRAVAOIALHCTALGRACAE 524
QY 503 CLMSRDPYCGMDOGRCSISY---SSERSVLOSINPAEPHECNPDPKAPLQKVSILAPNS 560
DB 525 CCLARDPYCAMDGSCTFPQAKRPRQDIRNGDPTLCSGSSHVLLEKLVGES 584
QY 561 RYLLSCPMESRHAATYSWR-----KENVDSCEPGHQSPPCLFENLTAQ 606
DB 585 GSAPLECEPRSLQAHVQWTFQAGAEAAHTQVLAERVERTARG-----LLRGLRRQ 636
QY 607 QYGHVFECAQGSYFREAOHQMLLPEDIGMAEHL 640
DB 637 DSGVLLCAVAEQGFQSLRLRLVHLVLSAAQERL 670

```


RA	Yee C.S., Chandrasekhar A., Halloran M.C., Shoji W., Warren J.T.,
RA	Kuwada Y.Y.: "Molecular cloning, expression, and activity of zebrafish semaphorin
RL	zla."
RT	Brain Res. Bull. 48:581-593(1999).
CC	- FUNCTION: MAY INFLUENCE OUTGROWTH BY A VARIETY OF GROWTH CONES
CC	INCLUDING THOSE OF THE POSTERIOR LATERAL LINE CANCILON.
CC	- SUBCELLULAR LOCATION: Secreted (By similarity).
CC	- DEVELOPMENTAL STAGE: EXPRESSED IN HIGHLY SPECIFIC PATTERNS WITHIN
CC	THE DEVELOPING EMBRYO.
CC	- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC	- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC	- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	or send an email to license@isb-sib.ch).
DR	EMBL; AF086761; AA043964.1;
DR	JFIN; ZDB-GENE-991209-3; sem3aa.
DR	InterPro; IPR003359; 1g_.
DR	InterPro; IPR003006; 1g_MHC.
DR	InterPro; IPR003659; Plexin-like.
DR	InterPro; IPR001627; Sema.
DR	Pfam; PF01403; Sema; 1.
DR	Pfam; PF00409; IG; 1.
DR	SMART; SMART0423; PST; 1.
RW	Signal; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.
KW	SIGNAL 1 17 SEMAPHORIN ZLA.
FT	CHAIN 18 860 SEMA.
FT	DOMAIN 241 539 SEMA.
FT	DOMAIN 645 724 IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN 722 858 ARG/LYS-RICH (BASIC).
FT	DISULFID 652 717 BY SIMILARITY.
FT	CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE 860 AA; 97263 MW; 5FD4C12194F5165C CRC64;
Query Match	17.2%; Score 621; DB 1; Length 860;
Best Local Similarity	26.5%; Pred. No. 4.3e-39;
Matches 194; Conservative 107; Mismatches 256; Indels 176; Gaps 31,	
OY	6 PGRAPSAPEARVPPPARLGIPLRLL-----LLMAAASAGH---LRSGR 53
F	17 PGRVAPOHTKENP-----RLKLSYNEMLESNNVTFTGLANSSGVDTPLMDGR 66
OY	54 IFAVWKGVADRDVDGQTEPTVLTFHPGGSSVWVGKGKYLLDFPEGKNASVRYNI 113
Db	67 -----GRLLVGAEDHVFSDDL-----VNI 85
OY	114 G-----STKGSCL-----DKRDCENVITLLERRSE-GLLAGTAKNRRHSWNL 155
Db	86 NRDVKQIAMPATPSKRDECKWAGKDRLRDCSNFEVNLISYNTHTYICGTGFPHIC-- 142
OY	156 VNGTVVPJGEM-----GAYAFSPDENSLVFEGDEVYSTRKQENYGK 199
Db	143 ---SLEMGKRAEDNIFRLDANYFENGKSKSYDPDKMOSSILLDGELYSIGSDAPMGKD 199
OY	200 IPFRIRIGESELTY--SDTVMPNDPEIKATIV-HODDAYDKITFYEFREDNPDKNEA 255
Db	200 FAIFETLGSNHPRIETEDHSRWLNPREFLGIHLIPESDNPEDDKTLFFPKERVAMGEHTG 259
OY	256 PLANSRVAQOLCRGDGGESSLSYSKWNFELKAMLVCSPAATN--KNNRRLDYVLL--P 310
Db	260 KATISRIOQLCKNDKGHRSL-VNKWTFLFKAKKLCSVPGELNGIDTFHEDELQVFLMSAK 318

```

OY 311 DPGSGWDPTRVYGV----SNPNVSAVCYUJSGDIDKFKRNSLKG-Y--HSLJPN----- 360
Db 319 DE---KPNVIAVTTSSNIFRGSATICMYSNADIRVF---LPIAHROGPNQWPF 370
OY 361 -----PRPGK-C-----LPDOPRIPTEFGVADRHBEVAORVEMPGKPLTPE---HS 404
Db 371 GCRVYPRPGTCPSKTCFGGFSTKDLRDPDVIEFARLHPAMYNPVQPMG--GKPIVPTNV 428
OY 405 KHYQKVAVHNRQASHGELFNHLYLTTRDGTIHNKYE--PGEENHFAFINM-EIOPFRRA 462
Db 429 EYQFOVLVVDREADDGQ-YDVMFEGTDLGVLKVVTTIPRESMOLDEEVLEEMTVFREP 487
OY 463 AAIQTMISDAEERKLYVSSQWESVOPRDLCEVYGGCGHGCMSRDPYCGMDOGCISIT 522
Db 468 TPITAMELSTYKQOULYLSGLDISQMPLRHCYVYKACAEGLADPYCAMDGTCSKVF 547
OY 523 --SSRSVYLOSTINAPAEHKECPNPKPD-----KAPLOKYSLANSRYIYSCPMESRHA 573
Db 548 PTKARRTRRDIROINDRDPLOSOSCDLHHNDLDEGYSVEERSYVGVNSSMFLTLSPKSORA 607
OY 574 TYSW-----RKEWVEOSCEPHOSPNCLITENTTAQVYHPCMAQEGSYFREA 624
Db 608 LIYWOLOPNRDKERKHEIYDERLSLTGGC-----LIRSLTQADSGVFLCHAVENHFIQPL 663
OY 625 Q--HWQLLPEDGI 635
Db 664 RRLNQVIPSQRV 676

RESULT 9
SK3B_HUMAN
ID SM3B_HUMAN STANDARD; PRT; 749 AA.
AC Q13214; Q93018;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3b precursor (Semaphorin V) (Sema V).
GN SEMA3B OR SEMA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer demarcation region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Dante M., Wamsley P.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
CC ENDOPLASMIC RETICULUM.
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
CC VARIETY OF NEURAL AND NONNEURAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL: U28369; AAD09138.1;
DR EMBL: U73167; AAC02731.1;
DR Genbank: HGNC:10724; SEMA3B.
DR MTM; 601281;
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR Signal; Immunoglobulin domain; Multigene family; Glycoprotein.
KW SIGNAL
FT CHAIN 1 24
FT SIGNAL 1 24
FT CHAIN 25 749
FT SIGNAL 25 749
FT CHAIN 239 537
FT SIGNAL 239 537
FT CHAIN 637 717
FT SIGNAL 637 717
FT CHAIN 699 702
FT SIGNAL 699 702
FT CHAIN 724 744
FT SIGNAL 724 744
FT DISULFID 644 710
FT CARBOHYD 82 82
FT CARBOHYD 124 124
FT CARBOHYD 427 427
FT CONFLICT 29 36
SQ SEQUENCE 749 AA; 83121 MW; 1F3B8F63F59444E3 CRC64;

Query Match 17.1%; Score 617; DB 1; Length 749;
Best Local Similarity 28.2%; Pred. No. 7.1e-39;
Matches 202; Conservative 92; Mismatches 293; Indels 130; Gaps 27;

QY 7 GRAAPAPRARVPPARIGLRRLRLLLMA-----AAAGQHRLRSGPRIFAVKQGHV 62
D 2 GRAAANV---IPG-----LALLMAVGLSAAAPRLRLSPFELQAMH----- 42
QY 63 GQDVNDGQTEPHVYLFEHPGSSVWVGKRGKYLFEDEPKNASKVRVNT----- 113
D 43 GLQTFSLERTCCYQALVDEGRGLFVGAENHVASLND---DNISKRKKLMPAPVEMR 99
QY 114 -----GSTKGSCLDRKRCENVYTLLEERSE-GLACGTNARHPSCNMLVNGTVPLG-- 164
D 100 EECWAGKADIT-----ECNMFVKLLNAVYNTNHLACGTGAPHPIC-----AFVEVGR 148
QY 165 -----EMNGAPSPDENSLVFEEDGVSTARKOYNGKIPFRIRGES 210
D 149 AEEPVLRLDPRIEDGKSKSPYDRHRRAASVLYGELTSGVAADLMGDTIFRSLGAP 208
QY 211 ELYT---SDTYMNPQETKA-TYHQDAQYDKTIYFFREDNPDKNP-APLNVSRVAQL 265
D 209 SLRTEPHDSRWLNPEPKFKVFWIPESENPDDDKIYFFRETAVEAAPALGRLSRVQCI 268
QY 266 CRGGGGGSSLSVSKMNTFLKAMLYCS--DAATKNENRLODVLPRPSGQWRTRYG 323
D 269 CRNVGGGGRSL-VNKMTEFLKARLVCVPGVEGDTHPDQDVL--SSRDHRTPLLYA 325
QY 324 VFSNP---WNTSAVCVYSGLDIDKYF-----RTSLKGHSSLNPPRGK----- 366
D 326 VFSSTSSIFGSGAVCYVSMNDVRAFLGPRFAHKEGPMQWMSYQGRVYPRRGKPSKTF 385
QY 367 --LPDQPIPTETQVADRHPVAVRPEPKPLTPLF--HSKYHYOKAVAHMQASHG 421
D 386 GTFSSSTKFPDDVIOFARNHPLMYSNVLPTG--GRPLFQVANYTFTQIADRVAAADG 443
QY 422 ETPIVIVITTRGTHKV--VEPGEQHSFAFNIMEIOPFRRAAIQIMSLDAERKIYV 479
D 444 H-YVLEFGTGVGLVYLVKIVPKGSRPSAEGLLLEELHVFDSAAVTSMTQISKRHLYV 502
QY 480 SSQMEVSGVPLDLCEVYGCGCHGLMSRDYPCGMDOGRGISYSS--BRSLQSTINPAP 537
D 503 ASRSVAVNOIALHRCAGHRCVTECCGLADPYCANDGVACTFPQSAKRFRQDVRRNDP 562
QY 538 HKCPNPKPKAPLD-QVSLAPNSRYIYLSCEMESRHATYSRH-----KENV 583
D 563 STLCSGSSRPALLEHKKVGVGSSAFLECEPRSLQARVEVTFQRAVYTAHTQVLAEBRT 622

```

```

QY 584 FOSCEPHGQSPNCILFIENLTAAQYGHFCEAOEGSYREAOHMLLPEDGIMAEHL 640
D 623 ERTARG-----LLRLRRRDSGVYLCVAVEGGFTQPLRLSLHVSATQABRL 671

RESULT 10
SM4B_MOUSE
ID SM4B_MOUSE STANDARD: PRT; 782 AA.
AC Q62179;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
DE SEM4B OR SEM4C OR SEMC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRI; TISSUE=Brain;
RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
RN [2]
RP INTERACTION WITH GIPC.
RX MEDLINE=99253973; PubMed=10318831;
RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT "A PDZ protein regulates the distribution of the transmembrane
RT semaphorin, M-SemF.";
RL J. Biol. Chem. 274:14137-14146(1999).
CC - FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC - SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC - DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC - SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC - SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X85992; CAA59984.1;
DR MGD: MGI:107559; Sema4b.
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR002165; Plexin-repeat.
DR InterPro: IPR001627; Sema.
DR Pfam: PF01437; PSI; 1.
DR Pfam: PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
KW Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT NO_TERM 1
FT DOMAIN 1 662
FT TRANSMEM 663 683
FT DOMAIN 684 782
FT DOMAIN 202 492
FT DOMAIN 548 608
FT DOMAIN 703 726
FT DISULFID 555 601
FT CARBOHYD 12 12
FT CARBOHYD 15 15

```

```

FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 574 574 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 782 AA; 86823 MW; 627A81FC8F87AC8 CRC64;

Query Match 17.1%; Score 617; DB 1; Length 782;
Best Local Similarity 30.4%; Pred. No. 7.6e-39;
Matches 193; Conservative 91; Mismatches 253; Indels 98; Gaps 28;

QY 75 HTVLFEHGGSSVWVGGRKYYLFD----FPEKMASVPTVNIIGTKGSL----DKR 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 16 YTALLLSQDCKTLVYGAREALFALNSNLSPFJGGEYDELLMSADADAKQCSFKGDKPKR 75

QY 125 DCEVYI-TLLERSEGLACGTNAHNSC-----NWLNV---GTVPLDGMGYAFS 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 76 DQWYIYITLLPLNSSHLTCTAASPICAYIHASFTLADDEAGNVI-LEDGKGHCPED 134

QY 174 PDENSLVFEDEYVSTIRKQEVNGKIPREFRIGESELYSDTV--WONQFIKATIVH 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 135 PNEKSTALVVDGELY-IGTVSSFGNDPAISRQSSRPTKTESLNLMDQDAFAVASATSP 193

QY 232 QDQAY----DGIYFFREDNPDKNPEAPLVNSRYAOLCRDQGESLSVSKNNTFLKA 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 194 ESLSPIGDDDKIYFFSEFGQEEFFENTIVSRVARCKDEGEERYLQ--QWMTSFLKA 252

QY 288 MLVCSDATNKNFNRLDQVFL--PDPSGQMDRTRYGVFSNPMVY-----SAYCVYSLGD 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 253 QILCSRDDGPEFNVLDVFTLNPNPO-DWRKTLISIGFTSQMHRGTGSAICVFTMND 311

QY 342 IDKYF-----RTSSLGKHYSLNPRRGKCLPD--QQPIETFTQVADR----- 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 312 VQKAFDGLYKKVNETQWYETHTQVPTPRGACITNSAREKRINSSQLDRLVNLFLKD 371

QY 384 ----HPEVAQRVERMGPRLKPLFHSKHYKQVAHVMQASGCFEHNLYLTTDGTITKV 439
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 HFLMDGVRSRLLLP-----RAYQRYAVHVRVGLH-STYVDVLTGTDGRLHKA 422

QY 440 VERPEQHSFAFNIMEIOPFRRAAIOITMSDAERRKLYVSSQMEVSQVPLDCEVGGG 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 423 VTLSSRHH-----TELOIFPGQPVONLLDSHGGLLYASHSGVQVVPANCSLY-PT 477

QY 500 CHGCLMSRDYCGMDGRC--ISIYS--SERVYLQISINPAEPHKECPNK-----PD 547
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 478 CGCDLLARDPYCAWTSACRLASLYQPLDASRPWTODIEGASVKELCNNSYKARFLVPG 537

QY 548 KAPLOKYSLAPNSRYLISCPRESRHATYSMRKENVESQCEPQGSPNC-ILFTENL--- 603
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 538 K-PEKVOAIOPTNTVTLACPLLSMLATRLWTH-----NGAPVNASASCRVLPGLLLV 590

QY 604 -TAQOYGHYFC-EAOGSSYFREAQHMOLLPEQIM 636
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 591 GSQOGLGVFOQCSIEBGFQOLVASYCEPVMEGVM 625

RESULT 11
SZ1B_BRARE
ID SZ1B_BRARE STANDARD; PRT; 778 AA.
AC 09M686;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Semaphorin 21B precursor (Semaphorin 1B) (Sema-21B).
GN SEMA21B OR SEMA3AB
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA MEDLINE=99425174; PubMed=10495275;
RX Roos M., Schacher M., Bernhardt R.R.;
RT "Zebrafish semaphorin 21b inhibits growing motor axons in vivo.";
RL Mech. Dev. 87:103-117(1999).
CC -1- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
CC VENTRALLY EXTENDING MOTOR AXONS.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF083382; AAD28103.1; -
DR ZFIN; ZDB-GENE-991209-6; sema3ab.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; plexin-like.
DR InterPro; IPR002165; plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01437; PSI; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1..17
FT CHAIN 18..778
FT DOMAIN 241..539
FT DOMAIN 645..723
FT DOMAIN 721..776
FT DISULFID 652..716
FT CARBOHYD 54..54
FT CARBOHYD 127..127
FT CARBOHYD 593..593
SQ SEQUENCE 778 AA; 88904 MW; 4D36F4323AE21895 CRC64;

Query Match 17.0%; Score 613; DB 1; Length 778;
Best Local Similarity 29.0%; Pred. No. 1.3e-38;
Matches 204; Conservative 111; Mismatches 260; Indels 128; Gaps 36;

QY 33 LLLLMAAASAOCHL-----RSGPRIFAVKGVH-GQDRVDF---GQTEPHVLFHEP 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 IVLLIWTLLIAFERKTVAKRSKSNVPRLKPSYKEMLESNLLTFNGLANSSAYHFFLDEE 66

QY 83 GSSSVWVGGRKYYLFDPEEKMASVPTVNIIG-----STKGSCL-----DKRCE 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 -RGRLEFGAKDHLVLD-----INMDQLISWSPSSPRDECKWAKGVOKKECA 117

QY 128 NYTLLERRSE-GLLACGTNAHNSCNLVNG-----TVPLGEM-----KTAAPSPD-E 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 NFIVYLPQFNQTHLYACGTGAFHVCAGHVEGKRSSEDTFRIGSSFENGKRSKYDRLQ 177

QY 177 NSVLVFEDEYVSTIRKQEVNGKIPREFRIGESELYT---SDTVMNQPIKATIV-HQ 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 TASMLIDG-ELIYAGTSADENMGKDFAIRFTLGKHHPIRTEQDSRWLNDPFRVSVHLIPES 236

QY 233 DOAVDDKIYFFREDNPDKNPEAPLVNSRYAOLCRDQGESLSVSKNNTFLKAMLVCS 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 DNAEDDKIYFFRENAIDGDISKATHARIGQLCKNDPFGGRSL-VAKWTFFLARLVCS 295

```


FT	CONFLICT	270	270	MISSING (1N REF. 2).
FT	CONFLICT	473	473	A -> S (1N REF. 2).
SO	SEQUENCE	785 AA;	88381 MW;	FE3FC796EEC1608E CRC64;
Query Match		16.5%;	Score 597;	DB 1; Length 785;
Best Local Similarity		26.0%;	Pred. No. 2.5e-37;	
Matches 194;	Conservative 116;	Mismatches 279;	Indels 156;	Gaps 32;
QY	35	LLMAAAS-----AQGLRSGPRLEPAWKGVHGQDRVDFG----	QTEPHYLFHEPG	83
Db	6	LLMAASLTGAMPSPHPDQDLHPATPRVRLSKELKATGTAFNPFLNTTQYRILLKED	65	
QY	84	SSSVWVGGRGVLYLFDEPEKKN-----ASPTVNIGSTKSGCLDRKDCENVTITLE-	134	
Db	66	HDRMWGSKDVLSTLDLHDINREPLIHMAASPORIECVLSGKDVN-GEGCNEVRLIQP	124	
QY	135	-RREGGLAGCTNARHPSC-----W-----NLNVTGVVP-----	162	
Db	125	WNRTH-LYVCGTGAIVNPCTIYVNGRRQAATPMTQTOAVRGSRGATGALRPMPTAPRQ	183	
QY	163	-----GEMKGVAFSPDENSLVLFEGDEYVSTIRKQENYGKIPRRRIRGESEL	212	
Db	184	DYIYLEPERLESCKGKCPYDPKIDLTASALINEELIYAAYIDPMGTDAIIRTLKQGTAM	243	
QY	213	YT---SDTWMONPQIKATIVHOD-QAYDQKIYFFREDNPKDKPEARLVNSRYAQLCRG	268	
Db	244	RTDYNSFWMLNDPSFIHAEILPDSAERNDDKLYFFERRSAAE-POSPAAYARIGRICLN	302	
QY	269	DQGGESSLSVSKWMTFLKAMLVCS---DAATNKNPNRLQDVFLLRDPDSGQARQTR---VY	322	
Db	303	DDGHCCCL-VKMWSTFLKARLVGSPGSDGJETHFDELQDVF-----QQTQDVANRPVYI	356	
QY	323	GVFSNP---WMYSAVCVYSLGDIDKVF-----TSSLKGYHSSLPNRPKSC----	366	
Db	357	AVFTSSSGSVFEGSNAVCIYSMADIRKVFGRPAHKEGPRYQWMPSPSGMKPIYRRPCTCGEQT	416	
QY	367	----LPDQPIPTETFOVADRHREVAQRVEREMGBLTPLF--HSHYUOKVAANRMQAS	419	
Db	417	FTPMKSTKDYPRDEYINPMRSHPLMYQAVUPL-QRRPLVYRTGAPYRLFTTIAVDQDA	474	
QY	420	HGEFFHYLYLTDDGSLTHKV-VERGEDEHSAFNIMELQPRRAALQIOTMSLDEERKLY	478	
Db	475	DGR-YEVLEFETDGRQYQKVTLPKRDDELBELMLEEVEYFKDAPVYMTISSLKROOLY	533	
QY	479	VSSQMEVSQVYLDCEVYGGGCHGCLMSRDYQCGMDQRCISY---SSERSYLOSTINPA	535	
Db	534	VASAVGTYHLSLHCQAYGAACADCCAKRDYCAMD-GQACSRTTASSKRRSRQDDYRHG	592	
QY	536	EPHKESP--NKPDPKAPLOKYSL-APNSRYULSCPMESRNATYSW-----RHKP---	581	
Db	593	NPTRCGRFNSNAKMNAVESVOYGVAAPVLEQOPRSPQATYKWLPGRDGDRREIRA	652	
QY	582	-----NVQDSEFPGHOSNCLFTIENTLAQOYGHYFCFAQDGSYFFRQAHQOLLPRDQIM	636	
Db	653	EDRFLRTEQG-----LLRLAQLQSDRGYLSCTATENPNKHVVTRVOL-----	694	
QY	637	AEHLGHACALAAASLWGLVPLTTL	661	
Db	695	---HYLGRDAVHAA-----LFPPLSM	712	
RESULT 15				
SM22_BRARE				
ID	SM22_BRARE	STANDARD;	PRT;	764 AA.
AC	Q9W6G6;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Semaphorin 22 precursor (Semaphorin 2) (Sema-22).			
GN	SEMA22 OR SEMA2.			
OS	Brachydanio rerio (Zebrafish) (Danio rerio).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			

